The Distributed Video Tracking Testbed (DVTT)

A DARPA Seedling Project

Bob Bolles, Chris Connolly, Aaron Heller
bolles@ai.sri.com
(650) 859-4620

Feng Zhao, Jim Reich
zhao@parc.com
(650) 812-5078

March 27, 2003

Goal

Develop a shared resource
to support the development and characterization of distributed tracking techniques.

Approach

Develop, install, and document a flexible network of video nodes that can gather time-stamped uncompressed (or compressed) video data, gather "ground truth" data, track moving objects, and be accessible over the web.

Focused on developing the video nodes, network, and control interfaces, plus characterizing video tracking

Focused on developing distributed tracking algorithms that can work with video sensors, extending SensIT capabilities

UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

Concept

A Video Node

Price for basic node: ~$1600

Wireless

+

+

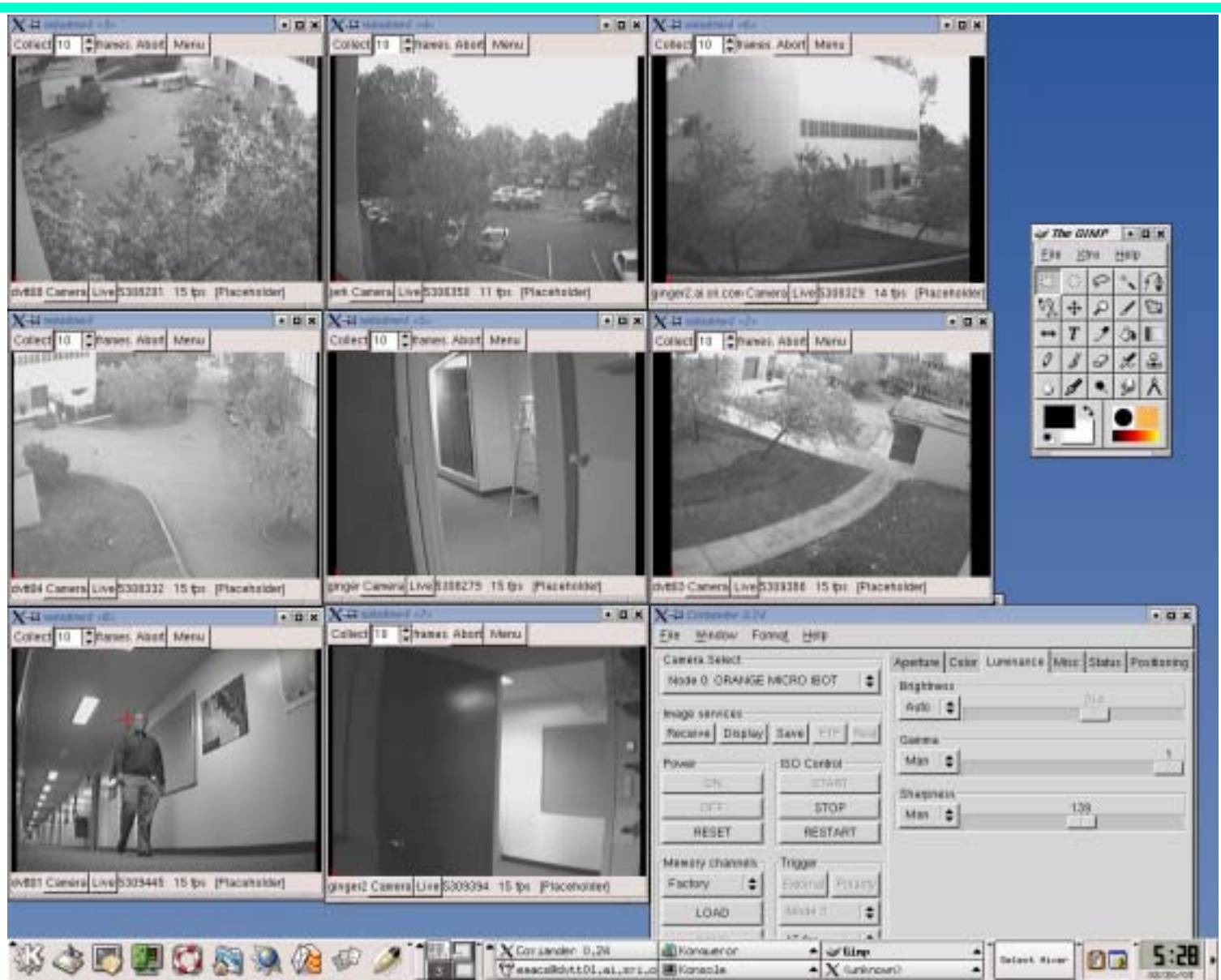
Gigabit Network

- **Network server**
- **Data acquisition & tracking software**
- **FREEdius computer vision software**
- **Firewire camera, DV, analog, or stereo**
- **TCP/IP networking interface**
- **LINUX operating system**
- **PC, Sun, SGI, or PC laptop**

Options

UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

Control Center Interface



Sample Data

SRI International

parc

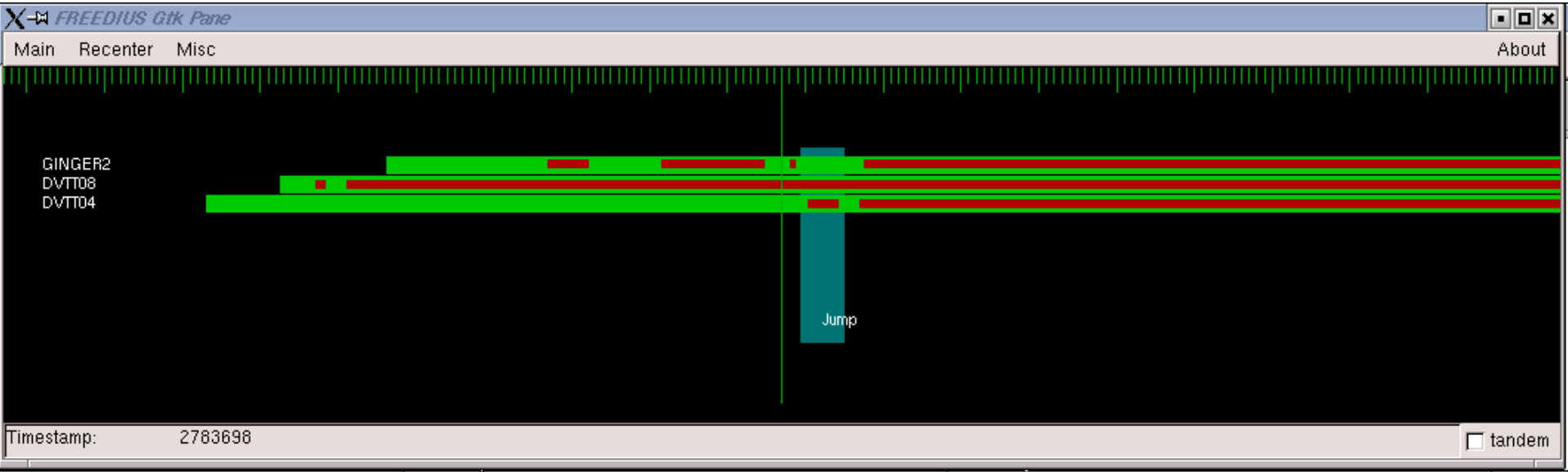
Video

UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

An Initial Browser for Multiple Time-Synchronized Videos and Events

SRI International

parc



Low-Light Sensing

SRI International

parc

Video

A Layered Tracker Using Motion Information

SRI International

parc

Video

UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

Initial Characterization of Video Tracking

Obersvations:

- When the tracker is “locked on” to a completely visible vehicle, the s.d. of x-y values for the centroid of the vehicle is approximately one pixel.
- The one-pixel s.d. is NOT a function of the range to the vehicle (or its size).
- The azimuth (and elevation) angle change caused by a one-pixel error is a function of the field of view of the camera. Most of our lenses cover 60 degrees horizontally. Therefore, an image that is 640 pixels wide, implies that a one-pixel error is only about 0.1 degree in azimuth.
- When a vehicle enters a dark shadow, the tracker often shrinks its rectangle, moving the CG by 15 or more pixels.
- The tracker could automatically identify “modes” that it is operating in, such as
  - something entering the camera’s field of view – don’t trust the CG yet
  - the vehicle is partially occluded by another moving object – don’t trust the CG as much
  - the vehicle is entering a known shadow – don’t trust the CG as much
  - the vehicle appears to be in the clear, use the 1-pixel-error estimate
  - …
- There are some systematic errors, such as a large shadow “attached” to the tracked vehicle, which offset the CG
- The CG estimate slides around when the vehicle turns
- There are many conditions that make it difficult to precisely locate a vehicle, such as rain, fog, low light (night), and a grey vehicle on a grey road.
- The current tracker has trouble if it never sees the whole object.

UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

PARC Cross-Node Video Tracking

Goal: Combine our SensIT-proven cross-node tracker with SRI's single-node Video tracker in the DVTT Testbed

- Cross Node Tracker
  - SensIT-like cross-node tracker collects and fuses measurements from multiple single-node trackers
  - Bayesian fusion approach supports multiple sensing modalities
- Sensor Tasking
  - Information-Optimal or Visibility-Based Sensor selection
- Multi-Target Tracking
  - Distributed Identity Management

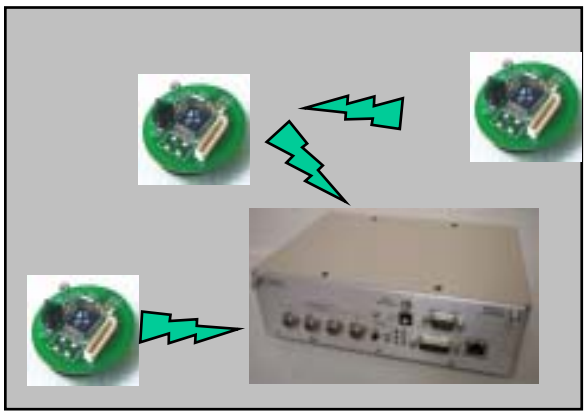
UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

PARC Collaborative Sensing Testbeds

SRI International

parc

Experiment at 29 Palms (in the Mojave Desert)
- **Various vehicle types**
- **Sensing modalities: acoustic, passive IR, seismic**



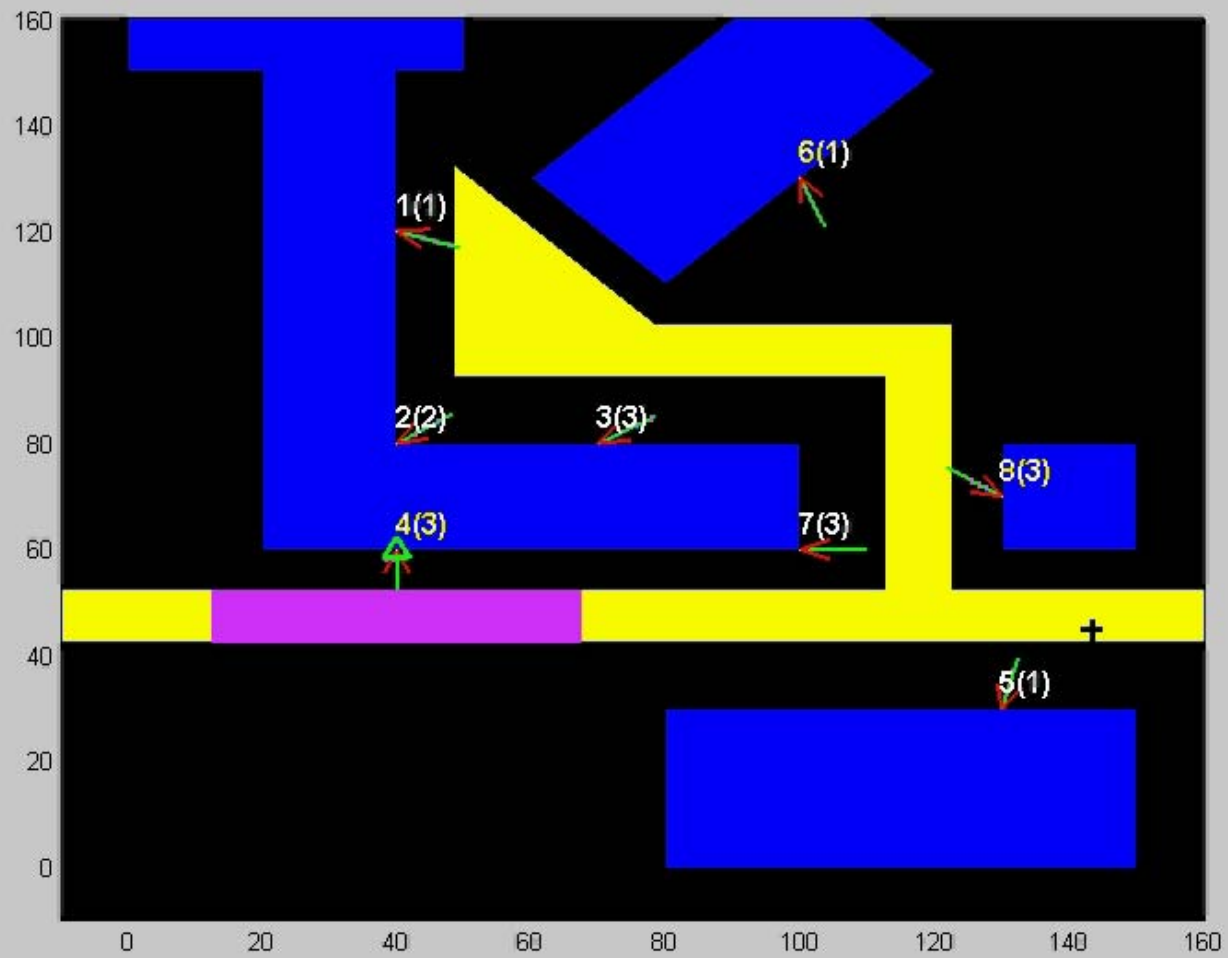
Ad hoc sensor networks: motes and sensoria nodes

100 Crossbow "Mica Mote" testbed

PARC Semi-Anechoic Chamber with 32-channel high-speed data acquisition

UNCLASSIFIED - Approved for Public

PARC Cross-Node Video Tracker



KEY

1 square=5m

- =Building
- =Road
- =Node
- =Leader Node
- =Active Node
- \+ = Target Location

Video

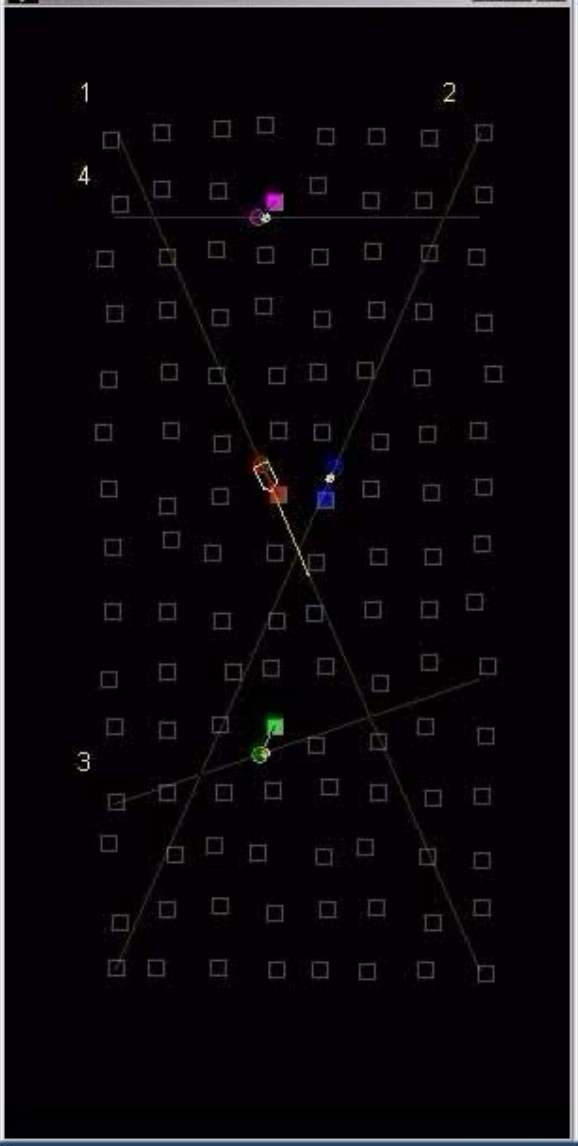
UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

Managing Multiple Target Identities in a Distributed Video Network

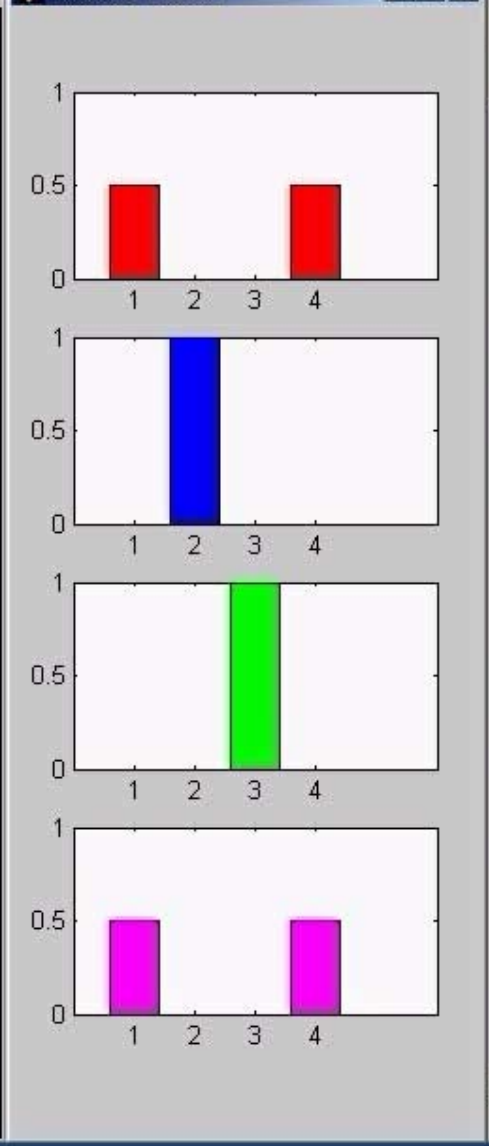
SRI International

parc

DASim



Belief of Leaders



MATLAB

File Edit View Web Window Help

Current Directory: C:\matlabR1

```
-------------------------------------------------------
-------------------------------------------------------
Node ID 49: 0  0  1  0
Node ID 54: 0,50009          0          0     0,49991
Node ID 59: 0,49991          0          0     0,50009
Node ID 85: 0  1  0  0
Belief sum: 1  1  1  1


group =

    [1x2 double]     [85]     [49]     [1x2 double]

Node ID 50: 0,15        0,15        0,55        0,15
Node ID 54: 0,79141     0,069531    0,069531    0,069531
Node ID 59: 0,1502      0,1502      0,1502      0,54941
Node ID 85: 0,15        0,54999     0,15        0,15

-------------------------------------------------------
-------------------------------------------------------
Node ID 85: 0  1  0  0
Node ID 50: 0  0  1  0
Node ID 54: 0,50009          0          0     0,49991
Node ID 59: 0,49991          0          0     0,50009
Belief sum: 1  1  1  1


group =

    [1x2 double]     [85]     [50]     [1x2 double]

Node ID 50: 0,15        0,15        0,55        0,15
Node ID 54: 0,79141     0,069531    0,069531    0,069531
Node ID 59: 0,1502      0,1502      0,1502      0,54941
Node ID 69: 0,15        0,55        0,15        0,15

-------------------------------------------------------
```

Ready

Video

DVTT "Products"

Now:

- Sample multi-camera data (in the SRIVIDEO format, a streaming video format that supports multiple time-synchronized video streams and several compression techniques)
- C code to read and write files in SRIVIDEO format
- C code to translate SRIVIDEO files into (and out of) several other standard formats, such as MPEG and AVI
- A multi-camera webcam-like interface that can be used from remote locations to help set up and run experiments using the DVTT
- An open source Lisp-C system call FREEdius that is a workbench for image and video processing with support for photogrammetry and object modeling. It is the open source successor to CME and 3Dius. It includes a browser for multi-camera SRIVIDEO formatted data.

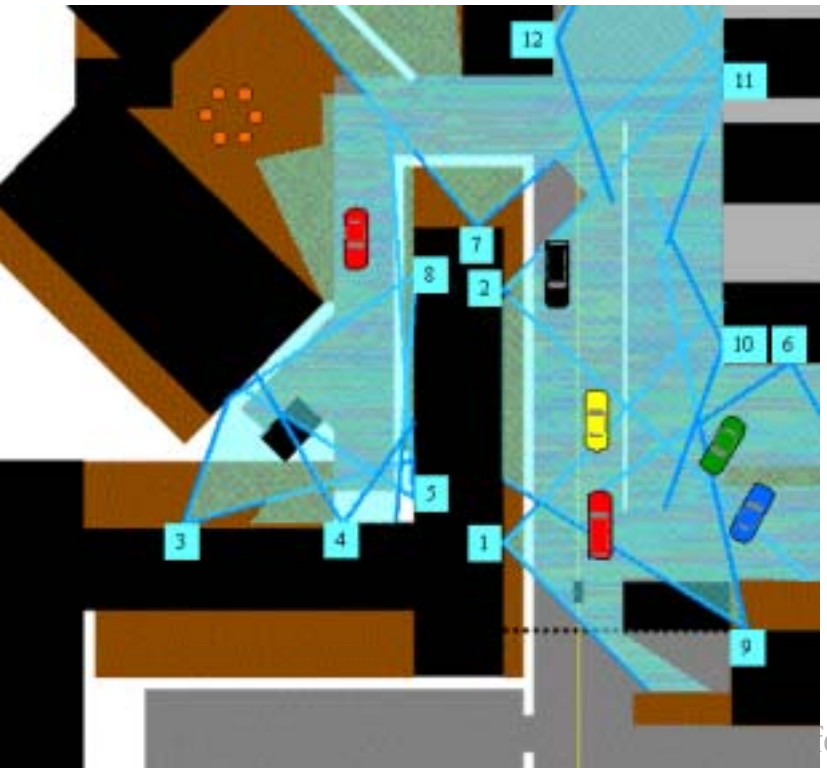
Soon:

- A report describing the lessons learned in developing the DVTT

UNCLASSIFIED: Approved for Public Release, Distribution Unlimited

Summary

Developed a Distributed Video Tracking Testbed
to support the development and characterization of distributed tracking techniques.



Plus

- A format to store and browse multi-band time-synchronized video streams
- An initial characterization of video trackers
- A cross-node video tracking technique (demonstrated in simulation)
- A technique for managing multiple target identities in a distributed network